1977

A Comparison of Mathematical Models of Multiple Cued Recall

Kathleen K. Biersdorff
Loyola University Chicago

A COMPARISON OF MATHEMATICAL MODELS OF MULTIPLE CUED RECALL

by

Kathleen K. Biersdorff

A Thesis Submitted to the Faculty of the Graduate School

of Loyola University of Chicago in Partial Fulfillment

of the Requirements for the Degree of

Master of Arts

July

1976

ACKNOWLEDGMENTS

The author wishes to express her deep appreciation to the members of her committee for all their helpful advice and prompt action. Especially appreciated is the continued support and encouragement of Frank Slaymaker, the Director of this thesis, and Deborah Holmes. A number of graduate students, notably Bernard Dugoni and Gary Hicks, have spent many hours patiently listening to and commenting on the ideas presented in this thesis. Their time and consideration is gratefully acknowledged.

VITA

The author, Kathleen K. Biersdorff, was born in Peoria, Illinois on September 29, 1953 and is the daughter of Mrs. Corinne (Molloy) Biersdorff.

She attended St. Boniface Grade School in Peoria graduating in 1966. She received a year's scholarship to the Academy of Our Lady High School in Peoria. While there, she received a Letter of Commendation from National Merit and was included in National Merit's Who's Who Among American High School Students. She graduated in the top ten percent of her class in 1970.

In September of that year she entered Loyola University of Chicago and received her Bachelor of Science, cum laude in June, 1974 with a major in psychology. While attending Loyola, she was a recipient of both Loyola University scholarships and Illinois State scholarships for four years. She was on the Dean's List during every semester of her attendance. In 1973 she was elected a member of Phi Sigma Tau, a national philosophy honor society. Between April, 1971 and June, 1974 she assisted Robert L. Solso of the psychology department in some of his research.

In September, 1974, she was accepted into the graduate program in Experimental Psychology and was granted an assistantship in psychology at Loyola University of Chicago with Frank L. Slaymaker.

To date, Ms. Biersdorff has two publications which are listed below.

Biersdorff, K., & Solso, R. A recodification of the Palermo and Jenkins word association norms, 1973.

Solso, R., & Biersdorff, K. Recall under conditions of cumulative cues, 1975.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

A vast amount of literature has been centered around the structure and processes of memory. A number of theorists and researchers (cf. Underwood, 1969; Nelson & Brooks, 1974) conceptualize memory in terms of attributes or distinctive features. Underwood (1969) has presented in detail some of these attributes. They include a temporal and a spatial attribute by which one determines when and where an event occurred, a frequency attribute by which one tells how often an event occurred, a modality attribute indicating which sense(s) experienced the event, and several attributes pertaining mainly to verbal or verbally mediated events; orthographic, associative non-verbal, and associative verbal attributes. The orthographic attribute describes the shape and spelling characteristics of the word in memory. The associative non-verbal attribute deals with acoustic, affective, and contextual information about the word. The associative verbal attribute places the word with respect to its taxonomic characteristics and a stable network of associations. These attributes are encoded with the word in order to fit the word into the existing structure of memory and to provide the necessary information to correctly decode the memory trace during retrieval. According to Underwood (1969), these attributes serve to "discriminate one memory from another and to act as retrieval mechanisms for a target memory."

Madigan (1974) has used differences between taxonomically cued and uncued recall to identify those aspects of attributes involved only in the retrieval processes. The cuing properties of the attributes are thought by many, including Madigan, to provide access to higher order units in which a target item is stored. These higher order units are the cognitive structure or organization that are imposed on individual items to facilitate storage and retrieval. This structure may be already existing in semantic memory, e.g., categories, or may be based on the contiguity of the items in a single episode. This structure must then be accessed in order to retrieve the individual items. The exact mechanism which performs this accessing is as yet undetermined.

There are several mathematical models currently in use which hope to explain this accessing process. This paper presents some of the theoretical foundations of the mathematical models and tests the predictions of these models in multiple cuing situations.

Bahrick's Generation-Recognition Model

According to Bahrick (1969), cues help the person generate possible responses from semantic memory and leaves the person with the task of recognizing one of the generated responses as the target. The probability of recall of the target in cued recall is the product of the probability of its implicit generation and the probability that it is recognized (Bahrick, 1970). The cues may be implicit, such as the recall of actual portions of targets found by Brown and McNeill (1966) with the "tip of the tongue" phenomenon. They state that subjects give themselves cues as to the identity of the target word by recalling features of the target, such as the first letter, the number of syllables, and the general sound of the word. As in the case of Bahrick (1969), the cues may also be explicitly presented by the experimenter. Furthermore, Bahrick (1969) says that cuing taps only the retrieval processes and not storage. This might imply that all attributes are represented automatically in the storage process since in order for storage of attributes to have an effect during retrieval, all attributes must be available for accession of the target memory. Or, this might imply that the question of storage of attributes is irrelevant to what happens in retrieval if the assumption is made that retrieval and storage are unrelated processes.

Tulving's Principle of Encoding Specificity

A somewhat opposing view of retrieval processes is expressed by Tulving and his associates in their principle of encoding specificity.

> It states that the properties of the memory trace of a word event are determined by specific encoding operations performed on the input stimuli; and that it is these properties, rather than the properties of the word in semantic memory, that determine the effectiveness of any given stimulus as a retrieval cue for the event. The principle suggests that if a stimulus in the retrieval environment renders possible or facilitates recall of the target word T, the retrieval information was appropriate to or compatible with the information contained in the episodic trace of T. Conversely, if a particular stimulus is ineffective in retrieving a particular trace, the conclusion follows that the appropriate relation was lacking. (Tulving, 1974, p. 778-779)

This suggests several things. First, what is stored determines what retrieval cues are effective in accessing memory. Second, the cuing task relies solely on the episodic properties of the target. Finally, there is an unbreakable bond between encoding and retrieval processes. While Tulving does not deny the possibility that there is a multiple encoding of attributes, he is careful in his own research not to make the assumption that any particular attribute is encoded.

In order to test the encoding specificity principle, Thomson and

Tulving (1970) presented a list of words to be remembered with a weakly associated input cue accompanying each word. This cue, during retrieval, allowed access to the representation of the item in memory presumably by increasing recall of the higher order units and of the contents of those higher order units. Cues that were not specifically encoded, even though normally more effective, were much less successful in providing access to the list word representation. It is possible that by providing an input cue and the instructions that the cue would help the subject remember the target, Thomson and Tulving directed the encoding of the word away from its more popular interpretation and into a more unusual interpretation. For instance, if CHAIR was the target and COMFORT was the input cue, an output cue of TABLE, which is normatively a higher associate of CHAIR than is COMFORT, would be relatively ineffective since the list representation of CHAIR would more likely be associated with the category of "things that are comfortable" than with the category "pieces of furniture".

Postman (1975a) has raised additional arguments against the generalizability of the encoding specificity principle. He suggested that Thomson and Tulving's results were at least partly due to expectations that the input cue would be made available to the subject at recall. These expectations were based on a series of set-inducing lists given to familiarize the subject with the procedure and to specifically maximize the probability that the target would be encoded in terms of the accompanying input cue. Postman modified Thomson and Tulving's technique by eliminating the set-inducing lists and by using a mixed list procedure with both strong and weak cues available for some of the words at input and output. With this modification, Postman found that although recall was in general higher when the cues were the same at input and output, "regardless of the condition of input cuing, strong output cues were substantially more effective than weak ones." Postman's (1975a) study presents some evidence that the extensive use of the weak input cues was primed by the use of set-inducing lists and that without this priming, the principle of encoding specificity fails to predict accurately.

However, despite the methodological questions raised for the Thomson and Tulving (1970) article, there is support nonetheless for the notion that cuing increases the recall of higher order units and their contents. Lewis (1974) presented lists of taxonomically related words in a categorically blocked or unblocked manner. Category names or category items not included in the lists were presented as cues. For both blocked and unblocked lists more items were recalled in the cued recall task than in the free recall task. This effect was mainly due to the increased accessing of the higher order units, in this case categories.

Cues themselves may act to either facilitate or hinder recall of a target. For cuing to have a positive effect, the cue must facilitate both the accessing of the higher order units <u>and</u> their contents or else facilitate one and have a neutral effect on the other. This point is supported by evidence that recall of category items may be reduced by using list items as cues for the remainder of the list (Roediger, 1973). This method facilitates access to the higher order units by providing some category information but disrupts recall of the contents of the higher order units by reducing the number of responses within the category that are still recallable. The organization of the list items at input is disrupted by the system of cuing.

Lauer (1974) presented a list of words either in random order or organized alphabetically or taxonomically. Then she gave output cues either consistent or inconsistent with the input organization. Recall was best when input organization and output cues were the same and taxonomic cues facilitated recall more than alphabetic cues. Therefore, in order to facilitate retrieval with cues, the organization imposed on the output by the type and ordering of the cues must be compatible with the input organization.

Cuing from the Associative Verbal Attribute

Probably the most discussed attribute in the cuing literature is the associative verbal attribute. This attribute is represented by taxonomic information and the type of associative information tapped in word association norms. Unlike orthographic and acoustic information, associative information about a word cannot be given as a single unit. A word may, and often does, have a number of meanings and the more frequently a word appears in the language the more meanings and associations the word is likely to have (Reder, Anderson, & Bjork, 1974).

A number of contemporary studies (e.g., Winograd & Conn, 1971) have looked at this multiplicity of meanings and associations in connection with specific encoding and retrieval cues. As was noted earlier, Thomson and Tulving (1970) found that normally high level associates were not effective retrieval cues when a different lower level associate was presented with the target word at encoding. They concluded that the high level associate was not encoded with the target. This conclusion has been given support by others (e.g., Tulving & Thomson, 1973; Tulving, 1974).

Light and Carter-Sobell (1970) presented homographs with an accompanying adjective which biased the encoding toward one semantic context or the other. Using a recognition task, they presented the noun only, the noun plus the input adjective, the noun plus an adjective biased toward the same semantic context, or the noun plus an adjective providing a different semantic context. They found better recognition for homographs with the same semantic context than with differing semantic contexts. Light and Carter-Sobell felt that recognition is similar to cued recall in which the cue is the phonological representation and concluded that "recall, unless it is cued, gives S no hint of which memory representations are appropriate ones to examine." Therefore, if a different meaning is cued at recall than the one that was encoded, the wrong memory representation may be found and the target will not be recognized.

Reder, Anderson, and Bjork (1974) gave further evidence that within the associative verbal attribute recall and recognition depend on the semantic interpretation. They assumed that, based on the fact that high frequency words have more associates than low frequency words, high frequency words also have more semantic interpretations and are more affected by changes in cues between encoding and retrieval. Reder *et al.* presented a list of high and low frequency words each with a weakly associated input cue. The subject was then asked to recall the targets given either the input cue, a strongly associated cue, or no cue. Since the input cue and strongly associated cue were more likely to tap the same semantic interpretation in the low frequency words than in the high frequency words, they expected that differences in recall with these types of cues would be greater for the high frequency words than for the low frequency words. They found that recall with the weakly associated input cue exceeded recall with the strong extralist cue by an average of 32.5% for high frequency words compared with 9% for low frequency words. They concluded that recall depends on recognition of the specific interpretation of the word originally encoded.

Tulving and Thomson (1973) had previously shown that words later recalled were not able to be recognized from among subject generated associations to strong associates not specifically encoded with the targets. In that study subjects, following encoding of target words and a weakly associated input cue, were required to generate associations to strong associates of the previously encoded targets and then circle any target words among the associations they had just generated. For example, if one of the targets was, in fact, TABLE, the subjects would be asked to generate associations to the word CHAIR (e.g., SOFA, TABLE, and SIT). Then the subjects would be asked to circle any of their responses that had been in the list they had seen earlier. If TABLE had been one of the generated associations and the subject circled it, the word was counted as having been recognized. While, in fact, the target items were often generated, they were seldom recognized using this procedure. However, Reder et al. (1974) in a second experiment, modified Tulving and Thomson's technique by using a 4-alternative forced choice recognition task and found that for low frequency words which had fewer semantic interpretations, recognition of targets from among the generated words was 84% as compared with 38% for high frequency words. They interpreted these findings to mean that for high frequency words the semantic interpretation given to the target within the generated words was more likely to be different from the original encoding than the semantic interpretation for low frequency words.

An issue of importance to the interpretation of the findings of Tulving and Thomson (1973) is the question of how the semantic similarity among the responses generated to the strong extralist cues affects the ability of the subject to discriminate between targets and distractor responses in the recognition task. Postman (1975b) suggested that in the generation task the subjects responded with words that are not only closely related to the strong associate of the target, but also closely related to each other. When asked to circle any target items from among the generated words, this semantic similarity would make the recognition task quite difficult. When a forced choice procedure was used in Tulving and Thomson's study, the number of items recognized more than doubled, suggesting a more relaxed decision criterion. On these grounds it might be suggested that the reason Tulving and Thomson found recall superior to recognition was that they had artificially forced the criterion for deciding that a generated item was from the list above the criterion used by the subject in the recall task.

Postman (1975a) was able to raise the raw recognition rate in yet another way. Instead of generating four or six associations for each word, his subjects generated only two associations. When only two associations were generated, recall was no longer superior to recognition of the targets from among the generated items. Two semantically similar responses left the subject with an easier recognition task than did four or six semantically similar responses, giving further support for Postman's assertion that semantic similarity between generated targets and distractors impedes recognition.

Tulving (1974) questioned the notion of semantic interpretation as an explanation of why normally high associates failed as retrieval cues. He used as input cues either the target itself, a low level associate unrelated to the high level associate used in his generation task, or a word that was congruous with the high level associate. For example, if the target was COAT, Tulving would use as his input cue either COAT, COVERING (a low level associate unrelated to the high level extralist cue LINING), or CLOTH (an associate of both the target and the extralist cue). This set of words congruous with the high level associates was generated by Tulving himself. Following list presentation, half the words were cued with high level associates and then for the other half of the list subjects were to generate four free association responses to the high level associates of the targets and then recognize any target words in their associations. In a second experiment the order of these two tasks was reversed. Finally, a cued recall task using the input cues was given. In the first experiment subjects performed much better on the extralist cuing task than on the recognition task but this effect was reversed in the second experiment, implying that there is an order effect. Also there was a statistically nonsignificant effect of extralist cues being more effective for congruous input cues over the other input cues. Recall with extralist cues was worst for words with no input cue.

In general, Tulving had difficulty explaining his results, especially the order of task effect. He concluded on the basis of the nonsignificant effect of congruous over incongruous input cues in the extralist cuing task that either the semantic interpretation of encoding is incorrect or else the "congruous encoding conditions were not congruous enough." This latter interpretation of Tulving's data is made all the more likely by his non-normative generation of cues congruous with the extralist cues.

The most surprising finding in Tulving's (1974) study is that extralist cues were least effective for those targets using the target word itself as the input cue. It was expected that the most frequently used meaning, which was supposedly tapped by the high level extralist cue, would be the meaning naturally encoded with the target (Light & Carter-Sobell, 1970; Winograd & Geis, 1974). Again, a methodological problem may provide the answer. Tulving (1974) suggests that the use of the target as cue may have induced subjects to encode the target solely in terms of its phonetic characteristics and because it was processed on such a low level, the semantic information provided by the extralist cue was less appropriate for these words. In any case, it is not clear that this result is typical of what happens normally in encoding and retrieval.

Multiple Cued Recall

Previously, encoding has been discussed in terms of single attributes, one at a time. Some of the findings discussed above seem to indicate that, within the associative verbal attribute, encodings may be limited to a single meaning at a time and that only cues congruous with that meaning are at all effective. Wickens (1970) suggested that words are encoded multiply and automatically. Underwood (1969) asserts that multiple attributes allow more than one path to a target memory. Nelson and Hill (1974) showed that the more opportunities to encode an item, the more encodings were made, and the better the stimulus was retrieved. Winograd and Geis (1974) used this same encoding variability principle in their study of homographs with equally probable or differentially probable interpretations and got results consistent with Underwood's suggestion. But neither of these studies specifically tested Underwood's assertion about multiplicity of retrieval paths across attributes.

One of the first studies to look at several attributes at once was Bregman (1968). Target words were presented serially with cues given within a continuing list. Cues were either graphic, phonetic, semantic, or a word contiguous with the target in the list. He found that the graphic cue was most effective when one, two, or three items intervened and the semantic cue was most effective when 24, 48, or 96 items intervened, implying that the amount of to be remembered material determines which cue is most efficient.

The present study was able to examine whether an associative, or semantic, cue has an advantage in effectiveness over a rhyme cue, which has both graphic and phonetic cue characteristics, at a delay comparable to that found in Bregman's 96-intervening items condition. There is a basic difference in the type of intervening material found in Bregman's task and the present experiment. While Bregman's subjects continued to view words and, presumably, organize them into some easily rememberable structure, in the present study only a limited number of to be remembered items follow (at most 14) and a mathematical task intervenes. This mathematical task is <u>per force</u> less verbal and may provide less specific interference than Bregman's continuing list. Also, with fewer items to be organized, the semantic characteristics may be less heavily relied on than with tasks involving many words and fewer good bases for organization.

There is, in fact, evidence for equal utilization of sensory and semantic characteristics as cues in recall of moderate length lists. Nelson and Brooks (1974) found no difference in cuing effectiveness of rhymes and synonyms when the pre-experimental probabilities of association and response set sizes had been equated. Both high and low levels of response probability were used for the rhyme and synonym cues. There were few intrusion errors at either level. If subjects were merely generating responses to the cues and giving the most probable response, then one would expect many intrusion errors for those cues with low probabilities of association. Their study was unable to distinguish between the possibility of the memory trace of a target being a single multifeatured representation or a number of single-featured representations because of the way their lists were constructed. They were able to establish, however, that sensory attributes can be functionally as important as semantic attributes in the representation(s). As such, Nelson and Brooks give support for the encoding specificity principle provided that multiple encoding takes place; and to the generation-recognition model, by virtue of the small number of intrusion errors in the low association responses. Clearly subjects compare the generated responses and recognize any targets.

While a word is encoded into a number of features each of which may be used singly to access the word in memory, often more than one feature at a time is used to reintegrate the target memory. Brown and McNeill (1966) in studying the "tip of the tongue" phenomenon gave some evidence that an almost complete recall of an item in memory consists of being able to recall various features of the item but either not being able to integrate them into a meaningful response or integrating them into an incorrect response that shares features with the correct response.

Using only the associative verbal attribute, McLeod, Williams, and Broadbent (1971) compared the effects of one and two retrieval cues on recall. Unlike Tulving and his associates, McLeod et al. gave no input cues with the list, thus avoiding the possibility of biasing the encoding toward one semantic encoding or another. Both associative cues were highly associated with the target response and were unassociated with each other. The target list was presented, followed by free recall. For each unrecalled word, first one cue was given and if still unrecalled, both cues were presented together. The two cues together facilitated recall much more than the single cue. McLeod et al. suggested that there might be an interaction between the two cues such that together they might produce a response, whereas the cues separately might not.

Independence of Cues in Multiple Cued Recall

Cues sometimes give information that is used independently of that given by other cues. Galbraith (1975) looked at the frequency attribute and a form of the associative verbal attribute to determine whether these two types of information could be used independently to make decisions about a target item. Pairs of words were designated correct or incorrect and presented one, two, four, or eight times within the list. Subjects were then asked to either select the most frequent of two pairs or state whether a pair had been designated correct or incorrect. Part of the time the subject's decision could be based on knowledge of information about either attribute, i.e., the correct pair also occurred more often than the incorrect pair. The rest of the time specific attribute information was required to make a correct decision. He found that subjects could, indeed, keep the two attributes separated in memory and make decisions concerning one attribute that were unaffected by the status of the other attribute.

Cues may function independently if the response universe, i.e., all of the possible responses given the constraints of the cues in terms of semantic and/or sensory characteristics, is the same for each cue. This is not to say that the likelihood that a particular response will be given or that the manner of accessing this set of responses is the same. Rather, the set of allowable responses defined by each of the cues is the same. In the case of certain attributes this has already been shown to be true.

Bahrick (1974) using a cuing paradigm looked at pictorial and position cues either separately or in combination. He presented a page of pictures and names in a "seating chart" format. After a free recall trial he gave the subject as a cue either the position on the page of one of the pictures, the picture itself, or the picture in the position in which it occurred on the page. Bahrick found that pictorial and position information were encoded independently. That is, it should be noted that in this experiment neither cue narrowed down the possible responses more than the other or both cues together, i.e., the response universe (those names involved in the experiment) is the same for both cues separately and together. While more definitive evidence regarding independence of cues should be based on comparisons of "joint and successive presentation" of cues as Bahrick asserts, it is at the same time necessary to consider the redundancy of information about the response universe provided by the cues.

There are instances in the cuing literature where combinations of cues are given which provide additional information about how the response universe is to be reduced. One of these is the McLeod et al. (1971) study discussed earlier. Both cues were associative but because they were not associated to each other, it may be assumed that they tapped slightly different semantic interpretations of the target. Therefore, each additional cue provided information about which elements of the response universe, given the previous cue, were no longer appropriate.

Solso and Biersdorff (1975) looked at recall with first letter, rhyme, and association cues. These three types of cues were used because of the lack of redundancy of information about the response universe provided by the cues, i.e., no information about the sound of the target or its initial letter was given in the association cue and vice versa. They presented a list of words followed by a free recall trial. Then either the first letter, a rhyme, or an associate of each of the unrecalled words was given. Combinations of the first cue and an additional cue were then given for the still unrecalled words and finally all three cues were presented. They found that the multiple cuing situations increased recall far beyond what would be expected if the cues acted independently. This result extended also to a set of control groups who did not see the list but were to deduce the target words from the cues alone. This would seem to imply that the action of the cues is predominantly in the generation phase. Solso and Biersdorff concluded that the multiple cuing restricts the number of implicit associative responses generated and enhances the probability that any particular generated response will be given. Furthermore, the efficacy of cues is inversely related to the number of associations it elicits and directly related to the probability of the target word in relation to the cue.

<u>The Present Experiment</u>

The present study provides an additional instance in which each added cue provides information about the changing state of the universe of acceptable responses, that is, each new cue puts additional constraints on the responses that may be given. A preliminary study was done to generate the rhyme cues to be used in the experiment. Based on strong normative associations, a number of rhymes both orthographically distinct from and orthographically similar to the possible targets were generated by the experimenter. A large sample of these rhymes was presented to a group of subjects with instructions to generate as many rhymes as possible in the 30 sec. allotted to each word. For the experiment only the first response to each word was tabulated. The rhyme cues for the experiment were then selected from those items for which the probability of generating the target as a rhyme was approximately equal to the probability of generating the target as an associate to the association cue.

Each subject in the experimental conditions was shown a target list of common words constructed using the results of the preliminary study. This was followed by a brief distractor task and a free recall trial. Each subject was cued for all of the target words whether previously retrieved or not. One group of subjects received rhyme cues while the other group received association cues. These rhyme cues and association cues had been previously equated with respect to probability of association to the target stimulus. Then those subjects previously receiving the rhyme cues were given the association cues, and vice versa. At this point subjects were uninformed as to which rhyme cue and which association cue were associated with the same word. Theoretically, these cues should act independently of one another. Finally, all subjects received both the rhyme and association cues together.

Two groups of control subjects saw and free recalled a list of words unrelated to the target list but equated to it in terms of word frequency. This was done to maintain procedural equivalence with the experimental groups. One group of control subjects was then required to generate a rhyming word for each rhyme cue given to the experimental groups while the other control group generated a verbal association to each association cue given to the experimental groups. Then those control subjects previously generating a rhyme response were required to generate a verbal association for each association cue and conversely those control subjects previously generating a response to the association cue were required to generate a rhyming word for each rhyme cue. Finally, both control groups generated a response to each of the combined rhyme and association cues seen by the experimental groups. This provided a guessing rate for the cues. A number of hypotheses were testable with this procedure.

The Hypotheses

Nelson and Brooks (1974) had used previously equated rhyme and synonym cues and found no difference in their cuing efficacy. They concluded that "cueing with either semantic or sensory attributes can provide equally effective access to the coded representations of target words primed in the context of a rapidly presented list of unrelated items." However, each cue was used as both a strong and weak cue for different words. This necessitated the construction of several target lists. On the other hand, other cuing studies (e.g., Bregman, 1968; Solso & Biersdorff, 1975) did not equate strength of association *a priori* but used a single target list and found definite order of cue efficacy effects. This experiment combined pre-equating of cues on strength of association and use of a single target list. It was expected that Nelson and Brooks (1974) would be supported in that there would be no difference in the effectiveness of the rhyme and association cues. This would imply that sensory and semantic attributes can be equally effective when encoding is not influenced by directing attention to specific characteristics of the items.

Biersdorff and Solso (in preparation) presented two recall cues for each target either at the same time or at different points in the cuing protocol. They found that there was little or no interaction when the cues were presented separately. However, because subjects were given no more cues once the target was retrieved, no rigorous measurement of the cues' independence in this type of task was possible. The present experiment was able to make more specific predictions about the effect of a first cue on the effectiveness of a second cue because it did not rely on the tacit assumption that once a target is retrieved by one cue, another cue will also produce the target solely due to the target having been previously retrieved. Rather, this study provided a direct test of that assumption. If recall for the cue when it is presented second is greater than when it is presented first, then it might be concluded that once a target is retrieved there are two memory representations, one corresponding to the item as retrieved and one corresponding to the item in the list representation, either of which may be retrieved in cued recall. While this issue of multiple representation is important to the study of retrieval processes in general, the mathematical models presented here take into account retrieval of either representation of the target and do not discriminate which of the representations is being retrieved. If, however, there is independent encoding of attributes as seems to be the case with semantic interpretations within the associative verbal attribute, then it would be expected that recall for the cue when presented first would be about the same as when presented second. It was hypothesized in the present study that independence of separately presented cues would in general be supported and that there would be little or no difference in recall between the cues when presented first and the same cues when presented second.

It has been suggested by Bregman (1968) that sensory and semantic cues are differentially effective based on the amount of time (or number of items) intervening between presentation of an item and cued recall of that item. It is possible that the reason that graphic cues are more effective at short intervals and semantic cues are more effective at long intervals is that the task being performed in the interval calls for organization of the material to be remembered on some meaningful and efficient dimension. At short intervals or with few intervening items a sensory-based code would be most efficient because extensive recoding of the target is not necessary. But as the amount of information to be remembered exceeds some limit, usually considered to be seven plus or minus two items, some higher order organization or chunking is necessary. A semantic recoding of information is more efficient in this instance. It would seem possible then that in Bergman's task items were initially coded in terms of their sensory features and as the need to retain these items for longer periods of time became evident they were recoded into some semantic structure. Therefore, cuing the item after a few items intervened would be more successful with a sensory based cue while cuing the item after a great number of items intervened would be more successful with a semantic cue.

In the present experiment there are fewer list items to intervene between presentation of any particular item and recall of that item. Also, the task which intervenes between the last list item presented and the free recall trial is non-verbal and should provide little associative interference. If the tasks should cause the sensory coding of the list to be interfered with more than the semantic coding, the rhyme cue would be less effective than the associative cue on the first cued recall trial and even less effective on the second cued recall trial since an associative cued recall task immediately preceded it. However, the present experiment is similar in terms of order and type of task to the Nelson and Brooks (1974) experiment, at least up to the second cued recall trial in which differential interference was not found. Therefore, it was predicted here that cued recall with a sensory based cue would not be interfered with more than cued recall with a semantic based cue and that there would be no difference in number of items recalled with rhyme or association on either the first or on the second cued recall trial.

Biersdorff and Solso (in preparation) also found that recall for those subjects who were shown the two cues together far exceeded recall for the two cues separately. But it was assumed that once the target was retrieved with one cue, it would be retrieved with that cue and an additional cue, an assumption somewhat more acceptable than that retrieval using one cue guaranteed retrieval using a different cue. It seems more likely that, since the same cue is presented in the two cue combination as was presented earlier alone, if the cue was effective alone, it should be effective in the combination. Continuing to cue targets that have already been retrieved was hypothesized to control for the possibility that the cue in combination with another cue might change the context enough so that the originally successful cue might not be effective. Furthermore, it was predicted that the suggestion of McLeod _et al_. (1971) that two cues insufficient to produce a response when presented separately might do so when presented together would be supported by these data. It was suggested that the mechanism responsible for this effect had its locus in the generation phase rather than the recognition phase or both phases together. As such, this effect would be found in the control groups which generated responses in, presumably, the same manner as the experimental groups but was not charged with a recognition task. Following a generation-recognition approach to cued recall, one would expect that both experimental and control groups when given two cues separately would generate a full set of responses for each cue while when given the two cues together would generate for serious consideration only those responses that fit the specifications of both cues. As such, there would be fewer responses to choose from when the two cues were presented together than when they were presented at separate times. The decision concerning which of the generated responses to give is based on chance in the control groups and chance plus recognition in the experimental groups. Chance, however, is a function of the number of items the response is to be chosen from and, therefore, is actually a function of the response generation phase.

This "chance" is a part of the experimental groups' response processes and is commonly referred to as guessing. It was therefore predicted that the effect of two cues together being more effective than the same two cues presented at separate times would be present in both the experimental and the control groups and that the effect was due to changes in the guessing rate, not cuing *per se*. Furthermore, because responding in the experimental groups was based on recognition in addition to chance, cued recall would exceed simple guessing of the target.

Finally, McLeod *et al.* (1971) presented several mathematical models in an attempt to predict response behavior in cuing situations. The predictions of these models, presented below, were compared for each level of cuing in this study to a mathematical model generated to fit Solso and Biersdorff's (1975) assertion that multiple cuing restricts the number of responses generated and enhances their probability of association.

Additive Model

The most widely used and simplest model for prediction of cued recall results has been the additive model which proposes that "there is no positive interaction between two cues in aiding retrieval" (McLeod *et al.*, 1971, p. 62). The probability that a first cue was successful in eliciting the target response in no way influences the probability that a second cue will elicit the target response. In the present experiment this model would predict that on the first cued recall trial the probability of success would be the probability that the target had been free recalled on the preceding trial plus the probability that the target could be guessed given the cue, i.e., the *a priori* probability of association between the cue and target, minus the product of the preceding two probabilities. Mathematically, this would be stated as

$$P(CR_1)=P(FR)+P(C_i)-P(FR)P(C_i)$$

where CR_1 = cued recall success with the first cue,

FR = successful retrieval of the target on the free recall trial, and

C_i = the a priori association between the cue and the target as measured with normative data or guessing rate.

As there is no interaction between the cues and since, in this experiment, both cues were approximately equiprobable, the probability of retrieving any target on the second cued recall trial would be a function, once again, of the probability that the target had been free recalled and the a priori probability of association between the cue and target. That is,

$$P(CR_2)=P(FR)+P(C_i)-P(FR)P(C_i)$$

Or $$P(CR_2)=P(CR_1).$$

On the final cued recall trial in which both cues are presented together, the additive model predicts that the probability of successful recall would be the probability of recall on the first cued recall trial, which includes the probability of free recalling the target, plus the a priori probability of association between the second cue and the target minus the product of these two probabilities, since they are independent. Mathematically, this becomes

$$P(CR_{1+2})=P(CR_1)+P(C_i)-P(CR_1)P(C_i)$$

Or $$P(CR_{1+2})=P(FR)+2P(C_i)-2P(FR)P(C_i)-\left[P(C_i)\right]^2+P(FR)\left[P(C_i)\right]^2.$$

Strength Model

The strength model, also presented by McLeod *et al*. (1971), "supposes that each possible response has a number attached to it, which is increased by the presentation of one cue, *and* by that of a second one" (p. 63). Each previous presentation of a cue increases the likelihood that the target response is given. On the free recall trial, successful retrieval is determined by the strength of the target response, as represented by a hypothetical "number", divided by the strength of the target and all other responses combined. This can be expressed mathematically as

$$P(FR)= \frac{R}{R+T}$$

where R = the strength of the target response on the free recall trial, and

T = the strength of all other responses combined.

Based on this equation, a further statement can be derived which will be, in the long run, computationally simpler.

$$1-P(FR)=1-\frac{R}{R+T}$$

$$= \frac{R+T}{T}$$

$$\text{And } \frac{P(FR)}{1-P(FR)} = \frac{R}{R+T} \times \frac{R+T}{T}$$

$$= \frac{R}{T}$$

On the first cued recall trial some constant amount of strength is multiplied with the free recall target strength. Assuming that the target is in the response set of the first cue, the probability of successful retrieval becomes

$$P(CR_1)=C\cdot P(FR)$$

where C = a cuing constant by which the original strength is multiplied to get the new strength for any cue.

Further,

$$\frac{P(CR_1)}{1-P(CR_1)} = C\cdot\frac{R}{T}$$

$$= C\cdot \frac{P(FR)}{1-P(FR)}$$

On the second cued recall trial, the strength obtained on the first cued recall trial is multiplied by this same constant such that

$$P(CR_2)=C\cdot P(CR_1)$$

And $$\frac{P(CR_2)}{1-P(CR_2)} = C^2\cdot\frac{R}{T}$$

$$= C\cdot\frac{CR}{T}$$

$$= C\cdot \frac{P(CR_1)}{1-P(CR_1)}$$

$$= \frac{P(CR_1)}{1-P(CR_1)} \cdot C\cdot \frac{\frac{R}{T}}{\frac{R}{T}}$$

$$= \frac{\left[\frac{P(CR_1)}{1-P(CR_1)}\right]^2}{\frac{R}{T}}$$

$$= \left[\frac{P(CR_1)}{1-P(CR_1)}\right]^2 \cdot \frac{T}{R}$$

$$= \left[\frac{P(CR_1)}{1-P(CR_1)}\right]^2 \cdot \left[\frac{1-P(FR)}{P(FR)}\right]$$

Converting this statement back in terms of the probability of cued recall success using the second retrieval cue, the strength model predicts that

$$P(CR_2) = \frac{\left[\dfrac{P(CR_1)}{1-P(CR_1)}\right]^2 \left[\dfrac{1-P(FR)}{P(FR)}\right]}{1+\left\{\left[\dfrac{P(CR_1)}{1-P(CR_1)}\right]^2 \cdot \left[\dfrac{1-P(FR)}{P(FR)}\right]\right\}}$$

On the final cued recall trial, the strength on the preceding trial is once again multiplied by the cuing constant such that

$$P(CR_{1+2}) = C \cdot P(CR_2)$$

And $$\frac{P(CR_{1+2})}{1-P(CR_{1+2})} = C^3 \cdot \frac{R}{T}$$

$$= C \cdot C^2 \cdot \frac{R}{T}$$

$$= C \cdot \frac{P(CR_2)}{1-P(CR_2)}$$

$$= C \cdot \frac{\frac{R}{T}}{\frac{R}{T}} \cdot \frac{P(CR_2)}{1-P(CR_2)}$$

$$= \frac{\left[\dfrac{P(CR_1)}{1-P(CR_1)}\right]}{\left[\dfrac{P(FR)}{1-P(FR)}\right]} \cdot \frac{\left[\dfrac{P(CR_1)}{1-P(CR_1)}\right]^2}{\left[\dfrac{P(FR)}{1-P(FR)}\right]}$$

$$= \frac{\left[\dfrac{P(CR_1)}{1-P(CR_1)}\right]^3}{\left[\dfrac{P(FR)}{1-P(FR)}\right]^2}$$

$$= \left[\frac{P(CR_1)}{1-P(CR_1)}\right]^3 \cdot \left[\frac{1-P(FR)}{P(FR)}\right]^2$$

Converting this statement back in terms of the probability of cued recall success using both retrieval cues, the strength model predicts that

$$P(CR_{1+2}) = \frac{\left[\frac{P(CR_1)}{1-P(CR_1)}\right]^3 \cdot \left[\frac{1-P(FR)}{P(FR)}\right]^2}{1+\left\{\left[\frac{P(CR_1)}{1-P(CR_1)}\right]^3 \cdot \left[\frac{1-P(FR)}{P(FR)}\right]^2\right\}}$$

Power Model

The power model is based on the notion that the response probability for each cue is solely dependent on the number of possible responses left. Each succeeding cue works to eliminate some proportion of the response universe. An additional assumption is made that each cue eliminates the same proportion of possible responses. On the free recall trial, the probability of successful retrieval for any one target is represented mathematically as

$$P(FR) = \frac{1}{S}$$

where S = the original number of responses in the universe.

On the first cued recall trial, and on subsequent recall trials, the response universe is reduced by some constant proportion. In other words, the cue delimits the responses that are still possible. Mathematically, this can be represented as

$$P(CR_1) = \frac{C}{S}$$

$$= C \cdot \frac{1}{S}$$

$$= C \cdot P(FR)$$

where C = the amount by which the cue divides the number of possible responses.

On the second cued recall trial, the response universe is again reduced by the same proportion. Thus

$$P(CR_2) = \frac{C^2}{S}$$

$$= C \cdot P(CR_1)$$

$$= C \cdot \frac{P(FR)}{P(FR)} \cdot P(CR_1)$$

$$= \frac{\left[P(CR_1)\right]^2}{P(FR)}$$

On the last cued recall trial the response universe is again reduced by the same proportion. Thus

$$P(CR_{1+2}) = \frac{C^3}{S}$$

$$= C \cdot P(CR_2)$$

$$= C \cdot \frac{P(FR)}{P(FR)} \cdot \frac{\left[P(CR_1)\right]^2}{P(FR)}$$

$$= \frac{\left[P(CR_1)\right]^3}{\left[P(FR)\right]^2}$$

Aggregate Model

This model combines features of each of the models presented above. Because the cues are first presented separately, independence of the cues is assumed for these trials. But when the two cues are presented together, the number of possible responses left in the response universe is reduced, as suggested by the power model and the strength of association between the cue and the target response increases as suggested by the strength model. This change as a whole follows the basic laws of conditional probability. As the response universe is reduced, the associated probabilities of the remaining responses assume a new denominator. For instance, if the target has a probability of response to both cue A and cue B of .40 and only one other word is associated with both cues but with a probability of .01, then when the two cues are presented together the probability of responding with the target is $\frac{.40}{.41}$ or .976 and the probability of responding with the non-target is $\frac{.01}{.41}$ or .024. The response universe has been reduced to two words with a combined probability of .41. As can be seen, this reduction of the response universe is of the order

$$P(CR) = \frac{R}{R+T}$$

where R = the probability of giving the target response, and

T = the combined probability of giving all other responses.

Clearly, this is the basic prediction of the strength model. While increasing the probability of association on each trial by a multiplier, this model actually is dividing the denominator, which is the response universe. It is a slightly more conservative form of the power model, which has been shown by McLeod et al. (1971) to overpredict response probabilities.

The aggregate model states that on the first two cued recall trials, in which the subject does not know which items of the two cue sets are for the same target, the additive model holds. On the trial in which the two cues are combined, the aggregate model proposes that there is a reduction of the response universe due to the constraints imposed by the first cue, the second cue, and the unique combination of the two cues. This prediction is most consistent with the strength model which is proposed to explain the data from the third cued recall trial.

Comparison of the Models

It is possible to distinguish the models on the basis of the predictions that each makes concerning the outcomes of the second and third cued recall trials. Because of the use of the free recall score and the first cued recall score as parameters in the power and strength models, it is impossible to derive independent predictions for these models on the free recall and first cued recall trials. However, because of the nature of the models it can be safely predicted that the strength that is added and the number of responses by which the universe is reduced give the same probability of response as the additive and aggregate models. A graphic representation of the predictions of each model for each trial in the experiment is presented in Figure 1.

Assuming that the free recall score is some arbitrary x where x is a value between 0 and 1, the additive model predicts that the first cued recall score in which the *a priori* probability of association between cue and target is .20 would be $x + .20 - .2x$ or $.8x + .2$. As stated above, at this level all of the models make the same prediction. At the second level of cued recall the predictions of the models differ. While the additive and aggregate models predict that the cues are independent and that the response probability is the same as on the first cued recall trial, the strength model predicts that the response probability would be $\dfrac{\left[\dfrac{.8x+.2}{1-(.8x+.2)}\right]^2 \cdot \left[\dfrac{1-x}{x}\right]}{1+\left\{\left[\dfrac{.8x+.2}{1-(.8x+.2)}\right]^2 \cdot \left[\dfrac{1-x}{x}\right]\right\}}$, and the power model

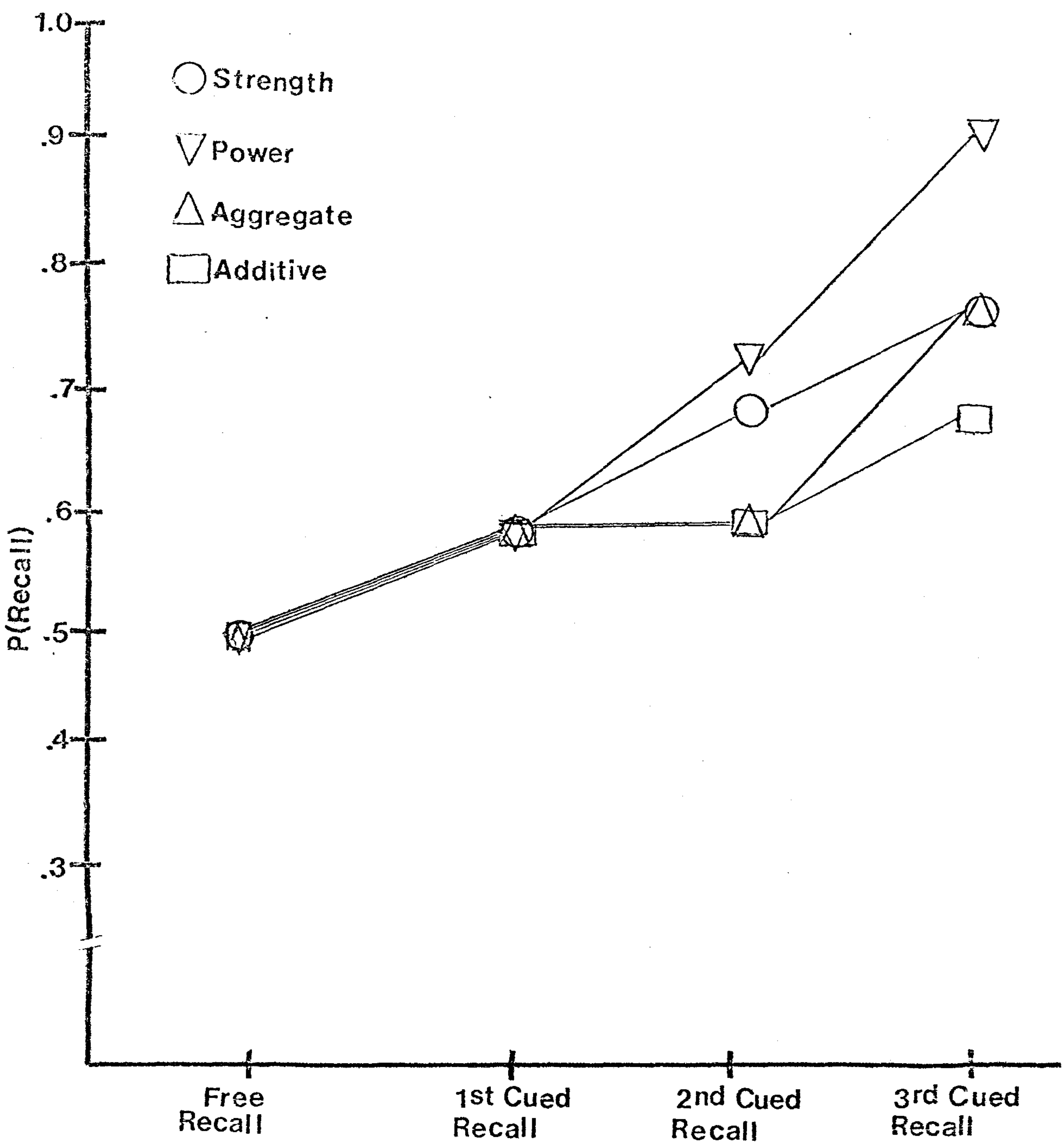

Figure 1. A comparison of the models in terms of their predictions on each trial (using .5 as the probability that an item is free recalled.)

predicts that the response probability would be $\frac{(.8x+.2)^2}{x}$ or $.64x + .32 + \frac{.04}{x}$. On the third cued recall trial in which both cues are presented together the additive model predicts that the response probability would be $.8x+.2+.2-.2(.8x+.2)$ or $.64x+.36$; the strength and aggregate models predict a response probability of

$$\frac{\left[\frac{.8x+.2}{1-(.8x+.2)}\right]^3 \cdot \left[\frac{(1-x)}{x}\right]^2}{1+\left\{\left[\frac{.8x+.2}{1-(.8x+.2)}\right]^3 \cdot \left[\frac{(1-x)}{x}\right]^2\right\}};$$

and the power model predicts a probability of response of $\frac{(.8x+.2)^3}{x^2}$ or $.512x+.384+\frac{.096}{x}+\frac{.008}{x^2}$. It was predicted that the aggregate model would be more accurate at all stages of the experiment on which the models can be distinguished because this model takes into account the nature of the interaction of the cues at each level.

CHAPTER II

METHOD

Preliminary Study

Subjects. One hundred fifty introductory psychology students fulfilling a course requirement served as subjects and gave responses for all list items.

Materials. Fifty target words were selected from the Palermo and Jenkins (1964) word association norms with the constraint that each target had a stimulus that elicited it with a probability between .15 and .25. The experimenter generated up to four common words that rhymed with the targets. No more than two of the four rhymes were orthographically similar to the target response. Since some of the targets had fewer than two rhymes that were either orthographically similar to or distinct from the target responses, the final list consisted of 151 stimulus words.

Each word was typed in capital letters on a standard 8½ x 11 page and was followed by two lines on which the subject was to record the rhyme responses. There were approximately 18 words per page and at least ten words were used with approximately the same number of subjects receiving each order.

Procedure. Subjects were given 30 sec. to write down as many words that rhymed with each stimulus as they could in the order that they thought of them. At the end of the 30 sec. a tone sounded and subjects were to move on to the next word. Subjects were instructed not to work ahead or to return to past words. While only first responses were tabulated, generation of multiple rhyme responses allowed the experimenter to measure the size of the response universe both in terms of the total number of words generated and in the number of first responses generated.

Main Experiment

Subjects. Eighty introductory psychology students fulfilling a course requirement served as subjects. Of these, 20 subjects were randomly assigned to each of the two experimental groups and the two control groups.

Materials. The target list was composed of 15 words with a frequency between 17 and 750 per million in the Kucera and Francis (1967) frequency count. None of the list words rhymed and there was little or no association among the words. The association cue for each target was taken from the Palermo and Jenkins (1964) word association norms as recoded by Biersdorff and Solso (1973) with a mean association probability of .196 and a standard deviation of .03. In addition, all targets were either the most popular or second most popular response for the association cues. The rhyme cues were selected from the stimuli presented in the preliminary study such that the average response probability for the target given the rhyme cue was .197 with a standard deviation of .032. All targets were either the first, second, or third most popular response to the rhyme cue. In addition, approximately half of the rhymes were orthographically similar (67% mean letter overlap) to the targets while the other half were orthographically distinct (44% mean letter overlap) from the targets. A priori probabilities for the combined cues were derived directly from the control group subjects who were to give responses to the cues in a free association type of task. The list of target words, rhyme and association cues, and their probabilities of association to the targets are included in Table 1. Five different orders of the target list were created so that each word appeared an equal number of times in each fifth of the list.

The equivalent list for the control groups consisted of 15 words chosen from the same level of the Kucera and Francis (1967) word count as the target list. In addition, none of the words rhymed with or was associated with any of the words in the target list as measured by Palermo and Jenkins (1964). These are also presented in Table 1.

Procedure. In the two experimental groups the target list was presented serially at a 2-sec. rate via a Lafayette memory drum. Five different list orders were used to control for primacy effects across subjects. Following presentation of the list, a 1-min. math computation task was given to control for recency. Subjects were then given one minute to write down as many list words as they could remember on a sheet of paper marked with 15 spaces, one per word. A single cue was then given for each of the target words regardless of whether the word had or had not been recalled. One experimental group received the rhyme cues and the other group the associative cues. Several orders of each set of cues were used with an equal number of subjects receiving each order in each group. Further, each cue appeared in each third of the list an equal number of times. Each cue was typed in capital letters on a 3 X 5 index card with its relation to the target in lower case directly below it. Each cue was presented for up to 15 sec. and the

TABLE 1

Targets, Their Rhymes and Association Cues with Accompanying Probabilities of Association, and the Equivalent Control List

TARGETS	RHYMES	ASSOCIATIONS	CONTROL LIST
Bed	Said (.19)	Sleep (.21)	Post
Heat	Sheet (.17)	Stove (.17)	Train
Arm	Farm (.25)	Hand (.15)	Jump
Girl	Whirl (.19)	Beautiful (.16)	Heart
Talk	Hawk (.21)	Speak (.25)	Born
Nurse	Worse (.15)	Doctor (.17)	Add
Home	Foam (.19)	House (.23)	Day
Church	Birch (.17)	Religion (.21)	Pan
Song	Wrong (.25)	Music (.16)	Rose
Hill	Bill (.22)	Mountain (.21)	Saw
Fast	Passed (.15)	Running (.24)	Short
Us	Bus (.17)	We (.19)	Terms
Here	Mere (.23)	Where (.19)	Growth
Soft	Coughed (.21)	Carpet (.21)	Nor
Lose	Whose (.21)	Find (.19)	Came

subject was required to make a response within this time, guessing if necessary. Responses were recorded on a sheet of paper with 15 spaces by each subject. No feedback was given as to whether a response was correct or incorrect.

Following presentation of the 15 cues, those subjects initially receiving the rhyme cues were given the associative cues for each target and those subjects initially receiving the associative cues were given the rhyme cues for each target word. The same general procedure was followed as with the first set of cues.

Finally, the subjects in each of the two experimental groups were shown both the rhyme and associative cues for each target word. Again, several orders of the cues were used across subjects. The two cues for each target were typed in capital letters side by side on a single index card with their relations to the target, i.e., association or rhyme, typed in lower case directly below. Rhymes and associative cues appeared in the left position an equal number of times. The same general procedure was followed as on the preceding two trials.

The two control groups served to measure the probability that subjects might generate the target response even if they had not seen the target list. Each of these subjects was shown the equivalent non-target list described above serially at a 2-sec. rate via a Lafayette memory drum. The same 1-min. math computation task that the experimental groups performed was then given. Following this, subjects wrote down as many of the words from the list as they could remember in 1 min. These free recall scores were compared to the free recall scores of the subjects in the experimental groups to assure that there were no initial differences in short term memory capacity between the experimental and control subjects. The two control groups were shown the same orders of cues and were required to give a word that rhymed with or was associated with the word on the index card, depending on the relationship shown on that card. The subjects were instructed that the list they had just seen was in no way connected with the words they were to generate. The control subjects were allowed 15 sec. to generate a response to each cue. If the subject had not responded in this amount of time, he was instructed to give a response quickly. Within these constraints, the control subjects, like the experimental subjects, generated responses first to one set of cues, then to the other set of cues, and finally to the two sets of cues together. A schematic of the design and procedure is presented in Table 2.

TABLE 2

Schematic of the Design and Procedure

	EXPERIMENTAL GROUPS		CONTROL GROUPS	
	GROUP 1	GROUP 2	GROUP 1	GROUP 2
LIST PRESENTED	Target	Target	Equivalent	Equivalent
	(One Minute Mathematical		Computation Task)	
FREE RECALL	Target List	Target List	Equivalent List	Equivalent List
FIRST CUE	Rhyme	Association	Rhyme	Association
SECOND CUE	Association	Rhyme	Association	Rhyme
THIRD CUE	Rhyme + Association	Rhyme + Association	Rhyme + Association	Rhyme + Association

CHAPTER III

RESULTS

As expected, the order of presentation of the cues (i.e., rhyme on the first cued recall trial and associative cue on the second cued recall trial or associative cue on the first cued recall trial and rhyme on the second cued recall trial) did not affect recall on either of the first cued recall trials. The experimental subjects recalled significantly more words ($p < .01$) than the control subjects. Mean recall scores for each group are presented in Table 3. This result may indicate that, while the control subjects were merely guessing some of the target responses, experimental subjects were able to rely on the presence of the target list in memory as well as the cues in generating their responses. Experimental subjects were doing more than deducing the correct response given the cues. The effects of cue order and previous experience with the target list on recall performance on the first and second cued recall trials were measured by a 2 X 2 X 2 nested ANOVA with repeated measures across the cues. The results of this analysis are presented in Table 4. This analysis indicated that the two cues presented separately are not independent. It had been hypothesized that if a word was initially recalled on the first cued recall trial, then it was no more likely to be recalled on the second cued recall trial with a different cue than if it had been unrecalled on the immediately preceding cued recall trial. As can be seen in Table 3, there was sig-

TABLE 3

Means and Standard Deviations of the Number of Target Words Recalled on Each Recall Trial by Each Group

	Experimental			Control		
Trial	RA Order	AR Order	$\overline{X}$ of AR & RA Orders	RA Order	AR Order	$\overline{X}$ of AR & RA Orders
Free Recall	7.5(2.12)	7.0(1.94)	7.25(2.0)	----*	----*	----*
1st cued Recall	8.6(2.17)	8.5(1.58)	8.55(1.84)	2.7(1.42)	3.4(1.43)	3.05(1.43)
2nd cued Recall	9.7(2.54)	9.7(2.58)	9.7(2.49)	4.8(2.49)	4.3(2.21)	4.55(2.31)
3rd cued Recall	11.9(3.28)	13.9(.99)	12.9(2.57)	10.4(2.95)	11.2(3.91)	10.8(3.40)

* Number of items recalled on the equivalent to target list.

TABLE 4

Summary of Analysis of Variance for First Two Cued Recall Trials

Source	SS	df	MS	F
Between Subjects	767.3875	9		
Order	.0125	1	.0125	<1
Exper./Contrl.	567.1125	1	567.1125	17.0006*
Order X E/C	.1125	1	.1125	<1
S(Order X E/C)	200.15	6	33.358	
Within Subject	158.5	70		
1st or 2nd cue	35.1125	1	35.1125	19.45**
1 or 2 X Order	1.5125	1	1.5125	<1
1 or 2 X E/C	.6125	1	.6125	<1
1 or 2 X Order X E/C	2.113	1	2.113	1.17
1 or 2 X S(Order X E/C)	119.1495	66	1.8053	
Total	925.8875	79		

* p <.01

** p <.001

nificantly better recall on the second cued recall trial than on the first cued recall trial ($p < .001$). A simple effects analysis measured whether this increase in recall between the first and second cued trials was present in both the experimental and control group protocols. The increase in recall across the first and second cued recall trials was significant for both the experimental subjects ($p < .01$) and the control subjects ($p < .001$) as seen in Table 4. A summary table of this simple effects analysis is presented in Table 5. This seems to indicate that once a person gives a particular response to a cue, that response is more likely to be given again when a cue is provided that generates that word as one of its associates or rhymes.

It was hypothesized earlier that cued recall with the rhyme cue would not be interfered with more than cued recall with the associative cue when preceded by a recall trial using the other cue type. A simple effects analysis measured the extent to which recall for a cue presented second exceeded recall for the same cue type when presented first. As can be seen in Table 3, there is a significant increase ($p < .05$) in recall across the first two cued recall trials for the rhyme cue and a marginally significant increase in recall ($p < .052$) across the first two cued recall trials for the association cue. The increase is approximately the same for both the rhyme and the association cues ($F=4.11$ for the rhyme cue and $F=3.81$ for the associative cue). The summary of this analysis is also presented in Table 5.

The probability that a word recalled or not recalled on a previous trial was recalled on a later trial is also measured using conditional probabilities. As can be seen in Table 6, the probability that a word recalled on the free recall trial is retained on the first cued recall

TABLE 5

Simple Effects in the 1st or 2nd Cued Recall Variable

Source	SS	df	MS	F
1st or 2nd Cue	35.1125	1	35.1125	19.4497***
1 or 2 at Experimental	13.225	1	13.225	7.326**
1 or 2 at Control	22.5	1	22.5	12.463***
1 or 2 X S(Order X E/C)	119.1496	66	2.113	
1 or 2 at Rhyme	18.225	1	18.225	4.1096*
1 or 2 at Association	16.9	1	16.9	3.8109
S(Order X E/C) + 1 or 2 X S(Order X E/C)	319.2995	72	4.4347	

* p <.05
** p <.01
*** p <.001

TABLE 6

Mean Conditional Probabilities of Recall

	Experimental		Control	
	RA Order	AR Order	RA Order	AR Order
$(CR_1 \mid FR)$	.353	.347	*	*
$(CR_1 \mid \overline{FR})$	.220	.207	*	*
$(\overline{CR_1} \mid FR)$	.140	.120	*	*
$(\overline{CR_1} \mid \overline{FR})$	.287	.327	*	*
$(CR_2 \mid CR_1)$	.440	.380	.073	.080
$(CR_2 \mid \overline{CR_1})$	.187	.267	.247	.207
$(\overline{CR_2} \mid CR_1)$	.133	.173	.107	.147
$(\overline{CR_2} \mid \overline{CR_1})$	.240	.180	.573	.567
$(CR_2 \mid CR_1 \cup FR)$	.527	.447	.073	.080
$(CR_2 \mid \overline{CR_1 \cup FR})$	.100	.200	.247	.207
$(\overline{CR_2} \mid CR_1 \cup FR)$	.193	.207	.107	.147
$(\overline{CR_2} \mid \overline{CR_1 \cup FR})$	.180	.147	.573	.567
$(CR_3 \mid CR_2 \cup CR_1 \cup FR)$	.707	.813	.377	.360
$(CR_3 \mid \overline{CR_2 \cup CR_1 \cup FR})$	.073	.073	.307	.377
$(\overline{CR_3} \mid CR_2 \cup CR_1 \cup FR)$	.133	.060	.067	.073
$(\overline{CR_3} \mid \overline{CR_2 \cup CR_1 \cup FR})$	.107	.053	.260	.180

* These probabilities are not reported as the free recall was made with a different list.

trial is approximately .35 regardless of whether that cue was a rhyme or an association cue. The probability that a word unrecalled on the free recall trial is recalled with a single cue is also approximately the same with the rhyme and the association cue. It is apparent from this table as well as the previous analyses that as the cued recall trials progress, recall performance is relatively stable in that both cues are about equally effective in cuing the targets regardless of whether the targets had or had not been retrieved on previous trials. This correspondence of conditional probabilities across the two cue orders is found for both the experimental and control subjects.

A 2 X 2 ANOVA measured the effect of having seen the target list and previous cue presentation order on recall on the final cued recall trial. A summary of this analysis is presented in Table 7. On this last cued recall trial the experimental subjects recalled a significantly greater number of target items than the control subjects ($p < .05$) and order of cue presentation on the previous cued recall trials (rhyme then associative cue or associative cue then rhyme) had no effect either alone or in interaction with the Experimental vs. Control variable.

As the number of cued recall trials increased, the recall scores of the experimental and control subjects became increasingly closer to each other. This effect can be seen in Figure 2 in which the mean probability of recall is presented for both cue orders of the experimental and control groups for each recall trial. It appears to be the case that as the recall performance of the experimental groups gets closer to unity the distance between the recall scores of the control subjects and experimental subjects becomes smaller. This effect may be due to the fact that the experimental subjects approached asymptote earlier in

TABLE 7

Summary of Analysis of Variance on the Final Cued Recall Trial

Source	SS	df	MS	F
Exper./Contrl.	44.1	1	44.1	4.93*
Order	19.6	1	19.6	2.19
E/C X Order	3.6	1	3.6	<1
S(E/C X Order)	321.8	36	8.938	
Total	389.1	39		

* $p < .05$

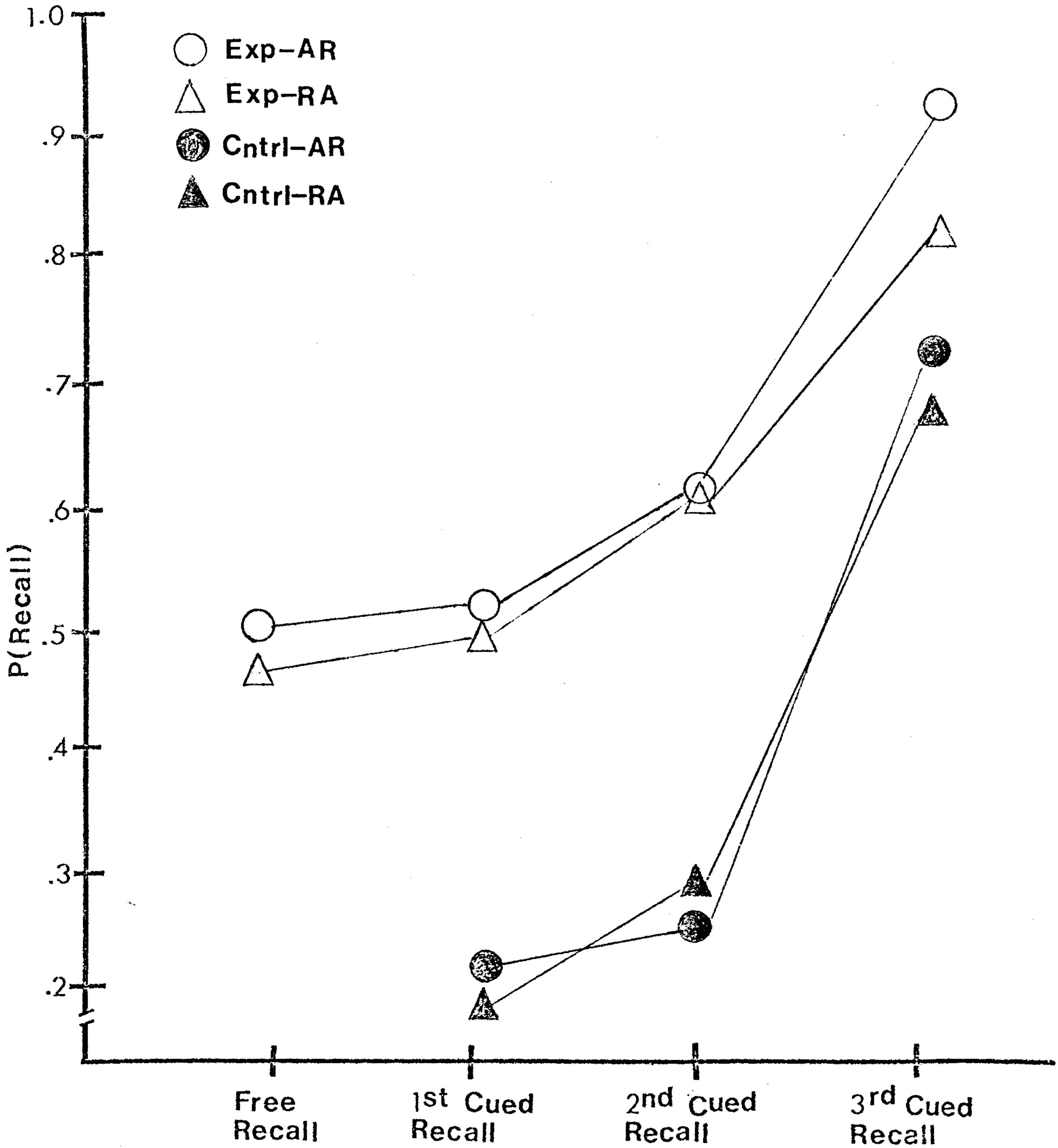

Figure 2. Mean probability of recall for the experimental and control groups on each recall trial.

recall performance than did the control subjects.

This earlier asymptotic performance of the experimental subjects may account for the results of two correlated t-tests of the hypothesis that the two simultaneous cues would produce better recall than the two cues separately presented. The measure of cumulative recall for the separate presentations was the total of the target items recalled on the free recall and the two cued recall trials. Within the control groups recall performance on the final cued recall trial was 10.8 words out of 15 possible while on the preceding trials combined 6.5 of the 15 possible words were recalled. The t-test performed on the control groups confirmed the hypothesis that the two cues together produced better recall than the two cues separately ($t=4.92$, $df = 19$, $p <.001$). However, within the experimental groups recall performance on the final cued recall trial was 12.9 out of 15 possible targets while on the preceding trials combined, 12.85 of the 15 possible targets were recalled. In the experimental groups recall performance with the two cues together was not significantly different from recall with the two cues separately ($t= .1365$, $df = 19$, $p >.10$). Several explanations for this finding are possible. It may be the case that recall performance on the free recall and first two cued recall trials is sufficiently high such that in order for performance on the final cued recall trial to be significantly better than the previous trials combined, last trial performance would have to be near perfect. (Last trial performance would have to be 13.61 out of 15 possible target words to be significantly better than the combined previous trials at the .05 level.) An alternative explanation for the difference in results of the t-tests is that the processes involved in cued recall are different from and do not overlap with the processes involved in deduction of the targets. This alternative seems less likely when the parallelism of the control group and experimental group recall curves is considered (see Figure 2).

Wilcoxon matched-pairs signed-ranks tests were performed individually comparing the predictions of each of the models with the data on both the second and third cued recall trials. The predictions of the models were determined by substituting the actual free recall score and the first cued recall score (or, in the case of the additive model, the a priori probability of association between cue and target) for the parameters of the model. Thus 20 predicted scores were obtained for each model. The means for the data, power, strength, and additive models on each trial are represented in Figure 3. The predictions of the aggregate model can be determined by connecting the prediction of the additive model in the second cued recall trial with the prediction of the strength model on the third cued recall trial.

One adjustment on the power model was made later. Since the power model predicted a probability of a correct response of greater than one on several occasions, and since better than perfect performance is not possible, all predictions of greater than unity were reduced to one. The revised predictions of the models are represented in Figure 4. Wilcoxon tests were performed on both the original and revised predictions of the power model.

On the second cued recall trial all of the models predicted the data well ($p < .05$ that the data do not differ from the predictions of the models) with the strength model slightly superior on this set of tests. However, the slight advantage of this model may be artificial since by summing the absolute differences between the prediction of

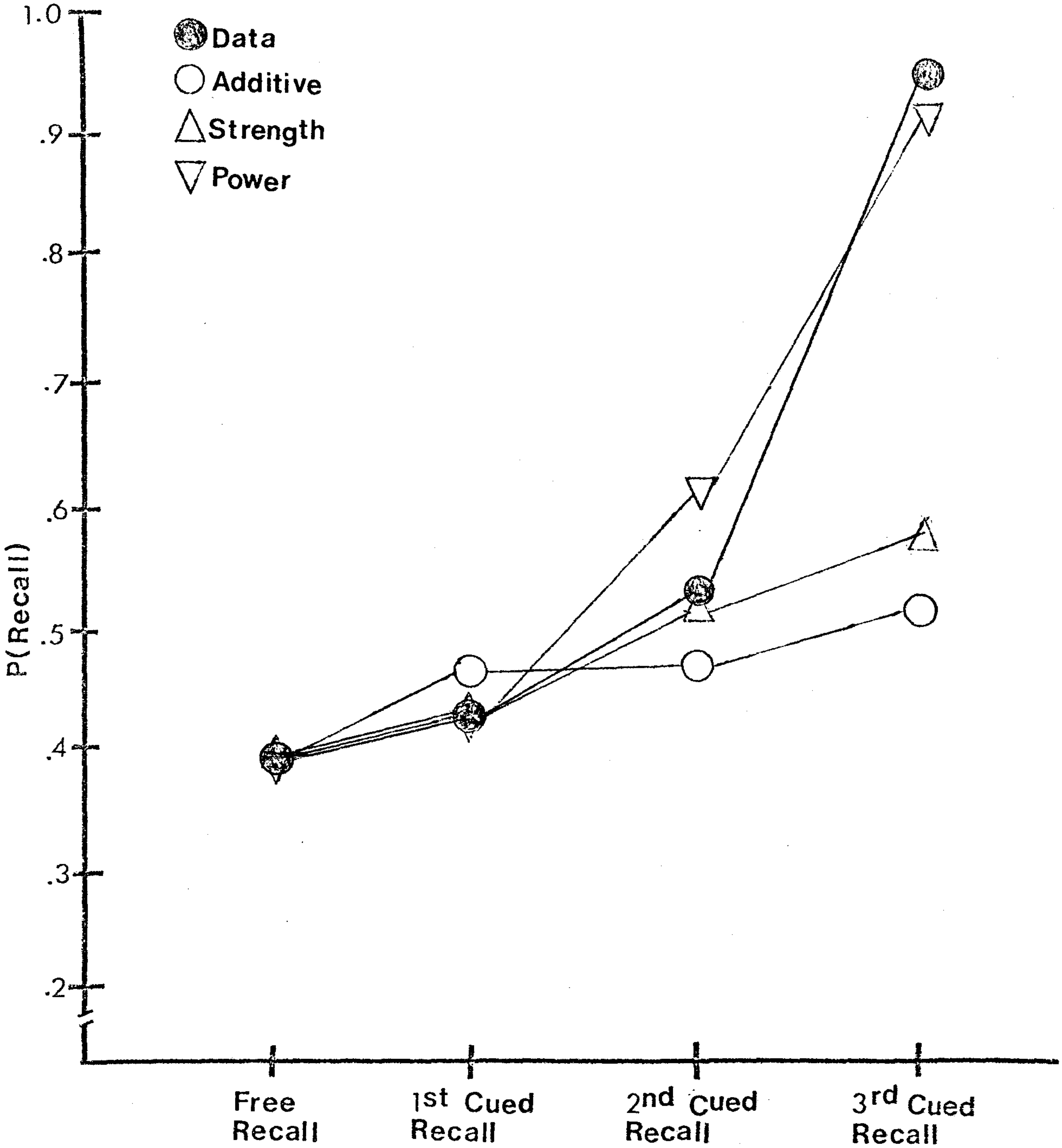

Figure 3. Mean probability of recall for the data and the models.

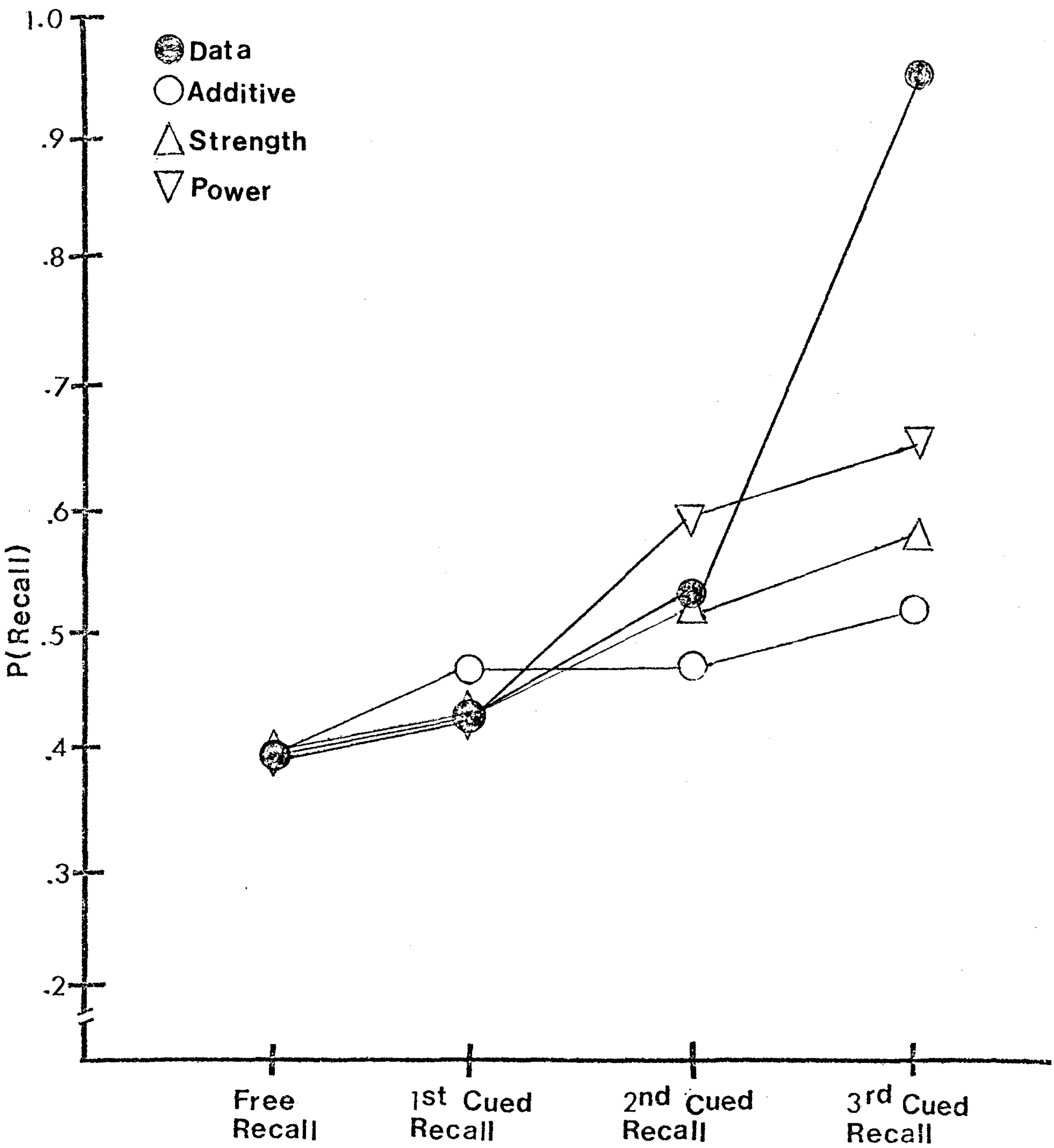

Figure 4. Mean probability of recall for the data and the models with the revised power model.

each model and the data it is trying to predict, the additive model is a slightly superior predictor than the strength model (mean absolute difference on a scale of 0 to 1 of .214 as compared with .226). Also, if a count is made of the number of times each model provides the best prediction of the three, the additive model once again has a slight advantage. (The additive model predicts best in nine cases, the strength model in seven cases, and the power model in four cases.)

On the final cued recall trial the Wilcoxon tests showed that the data exceeded the predictions of the strength and additive models ($p < .01$ for both). Using both the revised and unrevised predictions the power model predicts the data fairly well ($p > .05$). However, when a count is made of the number of times each model provides the best prediction of the three, the strength model predicts twice as well as either of the other two models. Furthermore, when the absolute differences between the predictions of the models and the data are summed and the mean absolute differences calculated, the additive model is a slightly better predictor than the other models. (The mean absolute difference on a scale from 0 to 1 for the additive model is .302, for the strength model .303, for the unrevised power model .459, and for the revised power model .317.) Clearly no firm statement can be made as to which of the models best fits the data. It depends on the criterion used. It appears, however, from looking at Figure 4, that none of the models predicts adequately in terms of the shape of the recall curve. Much of the difficulty stems from certain of the assumptions of the models not being met.

CHAPTER IV

DISCUSSION

A number of outcomes both predicted and unpredicted have been observed in this experiment. First, it seems clear from these data that cuing, in general, improves the number of target items that are retrieved since the recall scores of the experimental groups significantly surpass those of the control groups. This experiment, then, provides some evidence that cuing with a cue not explicitly encoded at input can aid in retrieval and that this improvement is not entirely due to guessing more targets.

Another finding of this study is that given equal a priori probabilities of association between the two cues and the target thay are associated with, both sensory-based and semantic-based cues can be equally effective within the context of a fairly short list. This is posited to be due to the fact that the shortness of the list did not necessitate that an extremely structured organization be applied to the list. It is quite possible that if a lengthy list had been used, the findings of Bregman (1968) would have been supported and the more highly structurable semantic organization would have led to a superiority of the association over the rhyme cue as an aid to recall. In the present experiment, however, it may be concluded that when no input cue is given which may bias the encoding in favor of one particular characteristic of the item, both sensory and semantic cues may be equally effective.

There also seems to be little or no proactive interference due to the nature of the first cue as witnessed by the facts that recall on the second cued trial exceeded recall on the first cued trial and that there was no order of cue effect on either the second or third cued recall trial. The data seemed to indicate that, contrary to the predictions made earlier, once an item is retrieved it tends to be more easily retrievable on later trials even when a different aspect of the item is cued. This is to be expected if one holds that more than one representation of the item now exists in memory for the successfully retrieved word or that the representation of that item in memory is somehow made stronger by each retrieval of the item. The implication of this finding, then, is that cues, even when presented separately, are not independent.

One finding of this study which moderates this conclusion is that guessing in the control groups followed the same pattern. There is a tendency to respond with the same word in situations which allow it. Since it was more likely to find a target list item which could fit both the rhyme and association cue constraints than a non-target item, there was a tendency in the control groups to repeat target items in later cued recall trials and to give different, possibly target, responses when the first response did not meet the constraints of the second cue. Therefore, the effect of non-independence of cues in the experimental groups could be entirely due to something inherent in the guessing strategy.

McLeod, Williams, and Broadbent (1971) had earlier considered the possibility that two cues presented together would aid retrieval more than the same two cues presented separately. This study provides evidence both contradictory to and supporting their contention. In an analysis of the number of different items recalled cumulatively over the first three recall trials as compared with the number of items recalled on the last recall trial, it was found that the experimental subjects did not recall significantly more items with the two cues together than with the two cues separately. Several explanations are possible. The most plausible of these is that the effect was confounded by asymptotic performance. Because approximately half of the items were free recalled and because each of the cues added somewhat to the number of items recalled, by the time the two cues were presented together, nearly all of the target words had been recalled on at least one of the previous trials. It should be noted that when a similar analysis was performed on the control data in which no free recall of the target list was possible, recall for the two cues together significantly exceeded the number of different items recalled on the preceding two single cue trials. In the case of the control groups, deduction of the target words with the two cues separately was not so great as to leave little room for improvement in recall when the two cues were presented together. If this is indeed what is happening, then the locus of the effect of recall for two cues together exceeding cumulative recall for the same two cues separately is in the response generation phase of retrieval.

The generation-recognition model of cued recall from memory serves as a good basis for differentiating the experimental and control groups according to the processes involved in choosing a response. While the experimental groups generate a response or a number of responses that fit the cues and then recognize one as the target (or guess), the control groups do not have a recognition phase. Therefore, anything which causes the experimental and control groups to have parallel recall performance is centered in the response generation phase or is random error. In the case of recall for two cues together exceeding recall the same cues separately, this is intuitively true as well. There are fewer responses that fit the constraints of the two cues together than fit either of the cues separately. Therefore, given non-asymptotic performance, it is to be expected, and it was found, that recall with two cues together exceeded cumulative recall for the same cues presented separately.

The Adjusted Model

It is clear from the results of the testing of the existing models that none of the proposed models as they stand provide a consistently good fit to the data. The power model comes closest. But when that model is not allowed to predict better than perfect recall, it assumes a shape quite different from the recall curve of the data.

The aggregate model fails in that it makes the assumption in the second cued recall trial that the cues act independently. The data, however, clearly point to the fact that retrieval on an earlier trial leads to a greater probability of retrieval of that item on a later trial no matter what cue is used. By removing the independence of cues assumption, a new adjusted model may be created which adequately fits the experimental group data no matter which criterion is used.

The free recall score once again serves as a parameter of the model. The probability of recall on the first cued recall trial takes into account that both the cue and retrieval on the free recall trial may raise the probability that the item is recalled. Mathematically, this is stated as

$$P(CR_1)=P(FR)+P(C_i)-P(FR)P(C_i)$$

where $P(C_i)$ = a priori probability of association between the cue and target, and is recognized as the prediction of the additive model.

On the second cued recall trial, recall may be due to use of the new cue or retrieval of items previously recalled on the free recall trial and the first cued recall trial. This takes the mathematical form

$$P(CR_2)=P(FR)+2P(C_i)-2P(FR)P(C_i)-\left[P(C_i)\right]^2+P(FR)\left[P(C_i)\right]^2.$$

Finally, the last cued recall takes into account the possibility that the target could be retrieved from any of the previous trials and takes the general form

$$P(CR_{1+2})=P(FR)+P(CR_1)+P(CR_2)-P(FR)P(CR_1)-P(FR)P(CR_2)-P(CR_1)P(CR_2)$$
$$+P(FR)P(CR_1)P(CR_2).$$

This formulation may be reduced to the two parameters of the free recall score and the a priori probability of association between the cue and target and thus becomes

$$P(CR_{1+2})=\left[P(FR)\right]^3+\left[P(C_i)\right]^3-\left[P(FR)\right]^3\left[P(C_i)\right]^3$$
$$+3\left\{P(FR)+P(C_i)+\left[P(FR)\right]^3\left[P(C_i)\right]^2+\left[P(FR)\right]^2\left[P(C_i)\right]^3\right.$$
$$\left.-\left[P(FR)\right]^2-\left[P(C_i)\right]^2-\left[P(FR)\right]^3P(C_i)-P(FR)\left[P(C_i)\right]^3\right\}$$
$$+9\left\{\left[P(FR)\right]^2P(C_i)+P(FR)\left[P(C_i)\right]^2-P(FR)P(C_i)\right.$$
$$\left.-\left[P(FR)\right]^2\left[P(C_i)\right]^2\right\}.$$

Predictions using the adjusted model were made on the basis of the free recall scores of the subjects in this study and the a priori probability of association between the cues and targets. The mean probability of the data and the predictions of the adjusted model are presented in Figure 5. The shapes of the two curves are nearly parallel and are quite close together in terms of their numerical values.

A set of Wilcoxon matched-pairs signed-ranks tests was performed on the first, second, and third cued recall trials with the adjusted model. In all cases the prediction of the model was not significantly different from the data ($p < .05$). Furthermore, when the absolute values of the differences between the data and the predictions of the model were summed, these differences proved to be smaller than between the data and any other model (2.795 and 2.1842 on the second and third cued recall trials respectively). By the criterion of best predictor among the models, the adjusted model was equal to the strength and additive models on the second cued recall trial and far superior to the other models on the third cued recall trial. (The adjusted model predicted best in twelve cases, the strength model in six cases, and the additive and power models in one case each.) In terms of all the criteria for model testing used in this study, the adjusted model was equal or superior to the other models in its ability to predict the data.

In summary, this study has presented evidence that cuing during retrieval can be effective by allowing the person to generate a number of responses consistent with the cue and leaving the person with the task of either guessing or recognizing one of the responses generated by the cue or from previously retrieved items. Further, a mathematical model consistent with the assumptions derived from the results of this experiment was built which fit the data according to a number of criteria.

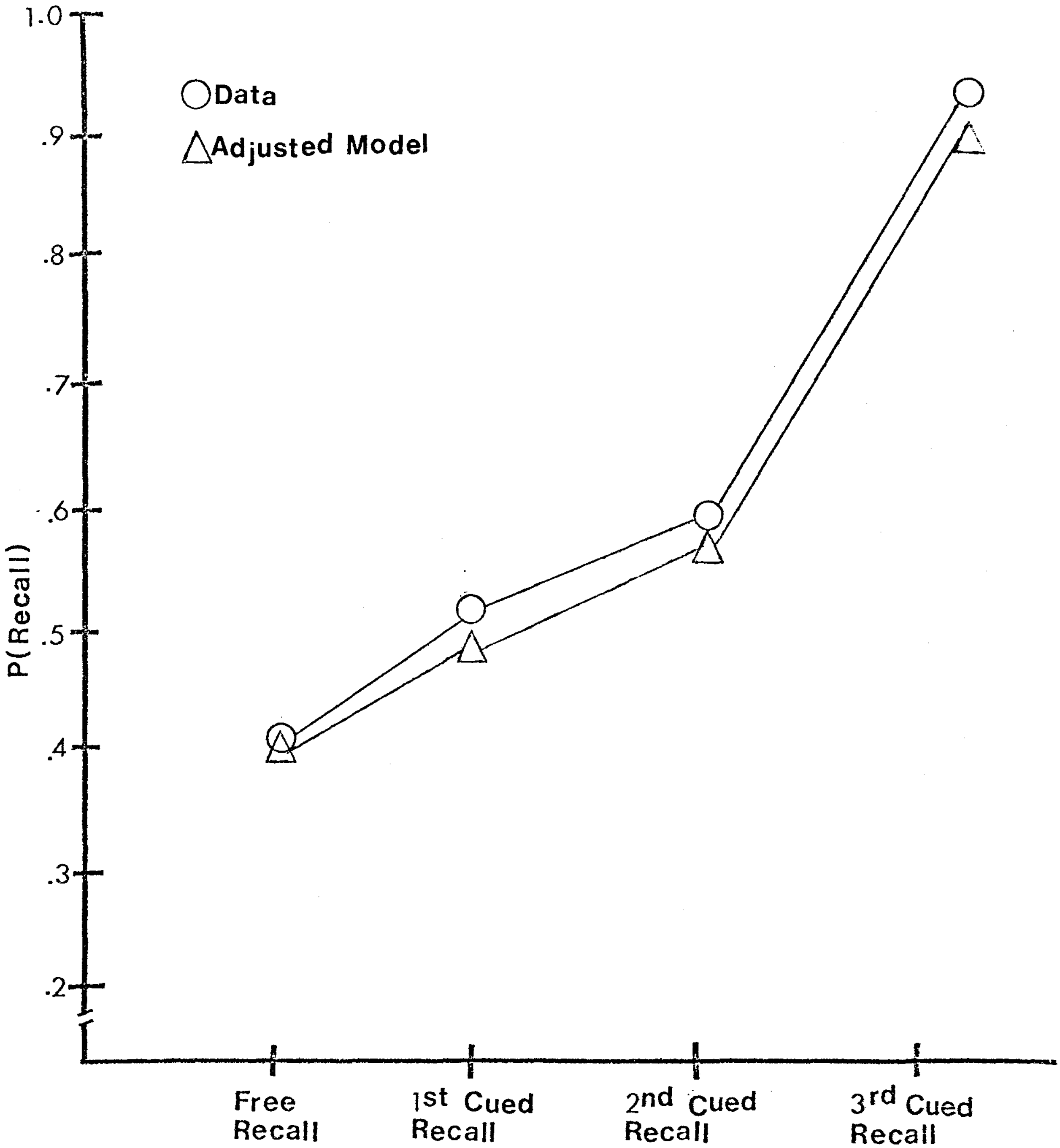

Figure 5. Mean probability of recall for the data and the predictions of the adjusted model.

References

Bahrick, H. Measurement of memory by prompted recall. Journal of Experimental Psychology, 1969, 79, 213-219.

Bahrick, H. A two-phase model for prompted recall. Psychological Review, 1970, 77, 215-222.

Bahrick, H. The anatomy of free recall. Memory and Cognition, 1974, 2, 484-490.

Biersdorff, K.K., & Solso, R.L. A recodification of the Palermo and Jenkins word-association norms. Behavior Research Methods and Instrumentation, 1973, 5, 14-22.

Biersdorff, K.K., & Solso, R.L. Cue interaction effects in multiple cued recall. (In preparation).

Bregman, A.S. Forgetting curves with semantic, phonetic, graphic, and contiguity cues. Journal of Experimental Psychology, 1968, 78, 539-546.

Brown, R., & McNeill, D. The "tip of the tongue" phenomenon. Journal of Verbal Learning and Verbal Behavior, 1966, 5, 325-337.

Galbraith, R.C. On the independence of attributes in memory. Journal of Experimental Psychology: Human Learning and Memory, 1975, 1, 23-30.

Kucera, H., & Francis, W.N. Computational Analysis of Present-Day American English. Providence, R.I.: Brown University Press, 1967.

Lauer, P. Encoding specificity in the cued and free recall of categorically and alphabetically organized words. Bulletin of the Psychonomic Society, 1974, 4, 496-498.

Lewis, M.Q. Cue effectiveness in cued recall. Journal of Experimental Psychology, 1974, 102, 737-739.

Light, L.L., & Carter-Sobell, L. Effects of changed semantic context on recognition memory. Journal of Verbal Learning and Verbal Behavior, 1970, 9, 1-11.

Madigan, S. Cue dependent effects in recall of categorized lists. Memory and Cognition, 1974, 2, 700-702.

McLeod, P.D., Williams, C.E., & Broadbent, D.E. Free recall with assistance from one and from two retrieval cues. British Journal of Psychology, 1971, 62, 59-65.

Nelson, D., & Brooks, D. Relative effectiveness of rhyme and synonyms as retrieval cues. Journal of Experimental Psychology, 1974, 102, 503-507.

Nelson, T.O., & Hill, C.C. Multiple retrieval paths and long-term retention. Journal of Experimental Psychology, 1974, 103, 185-187.

Palermo, D.S., & Jenkins, J.J. Word Association Norms. Minneapolis: University of Minnesota Press, 1964.

Postman, L. Tests of the generality of the principle of encoding specificity. Memory & Cognition, 1975a, 3, 663-672.

Postman, L. Verbal learning and memory. Annual Review of Psychology, 1975b, 26, 291-335.

Reder, L., Anderson, J.R., & Bjork, R. A semantic interpretation of encoding specificity. Journal of Experimental Psychology, 1974, 102, 648-656.

Roediger III, H.L. Inhibition in recall from cueing with recall targets. Journal of Verbal Learning and Verbal Behavior, 1973, 12, 644-657.

Solso, R.L., & Biersdorff, K.K. Recall under conditions of cumulative cues. Journal of General Psychology, 1975, 93, 233-246.

Thomson, D.M., & Tulving, E. Associative encoding and retrieval: Weak and strong cues. Journal of Experimental Psychology, 1970, 86, 255-262.

Tulving, E. Recall and recognition of semantically encoded words. Journal of Experimental Psychology, 1974, 102, 778-787.

Tulving, E., & Thomson, D.M. Encoding specificity and retrieval processes in episodic memory. Psychological Review, 1973, 80, 352-373.

Underwood, B.J. Attributes of memory. Psychological Review, 1969, 76, 559-573.

Wickens, D.D. Encoding categories of words: An empirical approach to meaning. Psychological Review, 1970, 77, 1-15.

Winograd, E., & Conn, C. Evidence from recognition memory for specific encoding of unmodified homographs. Journal of Verbal Learning and Verbal Behavior, 1971, 10, 702-706.

Winograd, E., & Geis, M. Semantic encoding and recognition memory: A test of encoding variability theory. Journal of Experimental Psychology, 1974, 102, 1061-1068.

APPROVAL SHEET

The thesis submitted by Kathleen K. Biersdorff has been read and approved by the following Committee:

Dr. Frank L. Slaymaker, Chairman
Assistant Professor, Psychology

Dr. Eugene B. Zechmeister
Associate Professor, Psychology

Dr. Mark S. Mayzner
Professor, Psychology

The final copies have been examined by the director of the thesis and the signature which appears below verifies the fact that any necessary changes have been incorporated and that the thesis is now given final approval by the Committee with reference to content and form.

The thesis is therefore accepted in partial fulfillment of the requirements for the degree of Master of Arts.

July 29, 1976
Date

Frank Slaymaker
Director's Signature